

УДК 004.853

Н.А. Соколова, П.А. Руденко, О.Т. Голуб

СИСТЕМА ТЕСТУВАННЯ ЗНАНЬ, ЯКА РЕАЛІЗУЄ ВІДКРИТУ ФОРМУ ТЕСТОВОГО ЗАВДАННЯ

Вступ

Важливим аспектом автоматизації процесу навчання є реалізація комп'ютерного тестування. Тестові завдання можуть надаватися тому, кого навчають оволодіти методами комп'ютерного тестування (далі: учень) у відкритій та закритій формах [1]. Комп'ютерний аналіз відповідей на завдання в закритій формі здійснюється досить просто. Одним з основних недоліків завдань у закритій формі є наявність "прихованої" підказки – вибирати відповідь набагато легше, ніж писати її повністю самостійно. Використання відкритої форми тестового завдання, коли учень вводить відповідь на природній мові, дає можливість гнучкіше здійснювати контроль знань, але автоматизувати такий контроль набагато складніше, ніж при використанні закритої форми. Комерційних систем, які реалізують відкриту форму тестування російською або українською мовами, не існує.

Дослідження автоматизації процесу тестування у відкритій формі ведеться в кількох напрямках. У праці [2] розроблено систему тестування, що ґрунтується на аналізі символічних послідовностей. Система доведена до практичної реалізації, але має ряд істотних недоліків. Вона дає змогу добре працювати з відповідями у вигляді формул та чисел, а з текстовими відповідями працює дуже неточно. Наприклад, при аналізі символічних послідовностей слова "вода" і "водка" мають високий коефіцієнт подібності. Семантичний аналіз відповіді як такий у цій системі не проводиться.

Другий напрям розробляється переважно професіоналами-лінгвістами. Так, у праці [3] докладно розроблено граматику опису еталонних відповідей, що враховує їх семантику, проте співставлення ведеться тільки з фіксованим набором еталонних відповідей за принципом повного збігу або незбігу.

Добрі результати досягнуто при використанні систем комп'ютерного перекладу. Відома система тестування у відкритій формі, що ґрун-

тується на системі машинного перекладу "Діалінг" [4]. У даній системі еталонні відповіді та відповіді учнів подаються у вигляді мережі фреймів. Далі ведеться порівняння мереж фреймів із врахуванням коефіцієнта подібності певних термінів. У даний час система "Діалінг" не доведена до завершення, тому використовувати її недоцільно. Крім того, ця система тестування не реагує на слова, які випадають із контексту речення. Аналогічних систем українського перекладу не існує. Тому розробка систем тестування з відкритою формою тестового завдання є актуальною.

Постановка задачі

Під тестовим завданням розумітимемо пряме, чітко сформульоване запитання, яке зумовлює однозначну коротку відповідь, причому запитання спільно з відповіддю є твердженням – істинним або помилковим, залежно від того, правильну або неправильну відповідь дав учень [5].

При створенні тестового завдання у відкритій формі викладач повинен сформулювати еталонну відповідь, що надається разом із завданням. Як наслідок маємо задачу співставлення двох текстів: еталонної відповіді та відповіді, введеної учнем.

Тому метою даної статті є створення системи тестування знань, орієнтованої на конкретну предметну область [6]. Основна вимога до системи полягає у створенні бази еталонних відповідей та бази можливих відповідей учнів на основі методології експертних систем. База можливих варіантів відповідей створюється в процесі дослідної експлуатації. При використанні системи остання може поповнюватися новими варіантами відповідей, тобто донавчатися. Співставлення еталонної відповіді та відповіді учня пропонується проводити методами нечіткої математики.

Формалізація опису еталонної відповіді

На кожне із запитань з курсу викладачем-експертом формується еталонна відповідь, яка в загальному випадку подається в такому формалізованому вигляді [7, 8]:

$$S_k[S_{k-1}(KS_{k-1}^{(1)}, KS_{k-1}^{(2)}), S_{k-2}(KS_{k-2}^{(1)(m)}, KS_{k-2}^{(2)}), \dots, \\ S_{k-n}[S_{k-n}^{(1)}(KS_{k-n}^{(1)(1)}, KS_{k-n}^{(1)(2)}), KS_{k-n}^{(2)}]], \quad (*)$$

де $S = \{S_1, S_2, \dots, S_k\}$ – множина семантичних відношень; $KS = \{KS_1^{(1)}, KS_1^{(2)}, KS_2^{(1)}, KS_2^{(2)}, \dots, KS_{k-n}^{(1)}, KS_{k-n}^{(2)}\}$ – множина лексем еталонної відповіді, причому кожній лексемі присвоюється параметр “важливість” у проміжку $[0, 1]$. Крім того, формується $SKS = \{SKS_1^{(r)}, SKS_2^{(r)}, SKS_3^{(r)}, \dots, SKS_n^{(r)}\}$ – множина можливих значень для лексем еталонної відповіді; $r = \overline{1, N}$, r – номер лексеми; N – загальна кількість лексем.

Таким чином, всю відповідь можна записати у вигляді множини семантичних відношень $S_p(KS_p^{(1)}, KS_p^{(2)})$, де $KS_p^{(1)}$ – перша ключова лексема, $KS_p^{(2)}$ – друга ключова лексема, S_p – семантичне відношення, рівень вкладеності відношень залежить від кількості лексем, які входять в еталонну відповідь.

Після виділення ключових лексем, які входять до еталонної відповіді, складається множина можливих варіантів значень лексем, що формують за сукупністю базу даних таких лексем, вказуючи коефіцієнти відповідностей даної лексеми еталонній, які належать відріzkу $[0, 1]$. База можливих варіантів значень поповнюється в процесі пробної експлуатації (навчання) системи, коли викладач задає ступінь відповідності лексем, застосованих учнем. Трапляються варіанти, коли ключова лексема складається з кількох підлексем (так звана складена лексема). Підлексема так само може бути складеною.

Розглянемо приклад. Запитання і еталонна відповідь: *Потік подій називається регулярним, якщо* події слідує одна за одною через однакові, строго фіксовані проміжки часу.

Дана еталонна відповідь розбивається на такі лексеми:

- 1) події;
- 2) слідує одна за одною;
- 3) однакові строго фіксовані;
- 4) проміжки часу.

Лексеми 2, 3 є складеними. У цьому випадку при формуванні формалізованого представлення еталонної відповіді складені лексеми записуються так: слідує за(слідує, одне за одним), і(однакові, строго фіксовані). Можливі варіанти із вказуванням ступеня відповідності еталонній лексемі даються як для підлексем, так і для складеної лексеми в цілому.

У процесі навчання системи формується база даних неприпустимих лексем [8]. Неприпустима лексема – це лексема, наявність якої у відповіді учня може анулювати як семантичне відношення $S_p(KS_p^{(1)}, KS_p^{(2)})$, в яке входить дана лексема, так і повністю всю відповідь.

Наведемо приклад. Відповідь учня: *Потік подій називається регулярним, якщо* події слідує одна за одною через випадкові проміжки часу.

Як неприпустима лексема для слова “моменти часу” тут виступає слово “випадкові”. Дана неприпустима лексема анулює всю відповідь.

Аналіз відповіді учня

Використання форми тестового завдання, при якій відповідь є продовженням запитання, дає можливість зменшити роль синтаксичного аналізу та зосередити увагу на семантичному аналізі. Пропонується обробляти відповідь, введену учнем, таким чином. Спочатку проводиться морфологічний аналіз відповіді. Мета такого аналізу полягає в отриманні основ (під основою слід розуміти словоформу з відсіченням закінченням) із значеннями граматичних категорій (наприклад, частина мови, рід, число, відмінок) для кожної з ключових лексем, тих, що надійшли на вхід системи [6]. На даному етапі відбувається виправлення граматичних помилок, які можуть траплятися у відповіді учня.

Далі, відповідно до опису еталонної відповіді і розробленої граматики, запропонованої в [7], здійснюється пошук лексем, що відповідають еталонним, із врахуванням семантичних відношень [7].

Розглянемо приклад декомпозиції еталонної відповіді на запитання, наведене вище:

де $[1-5]$ – номери семантичних відношень, якими зв’язані ключові лексеми або інші семантичні відношення, що входять до еталонної відповіді. Відповідь учня співставляється з даним формалізованим поданням еталонної відповіді. Необхідно відзначити, що можливі різні варіанти відповіді в цілому. Наприклад, еталонне запитання спільно з відповіддю “*Потік називається стаціонарним, якщо* ймовірність $P_k(t)$ того, що за відрізок часу t виникне K подій” залежить тільки від його довжини t і не залежить від його розміщення на часовій осі.

Другий варіант відповіді “*Потік називається стаціонарним, якщо ймовірнісні характеристики стали в часі*” так само буде допустимим.

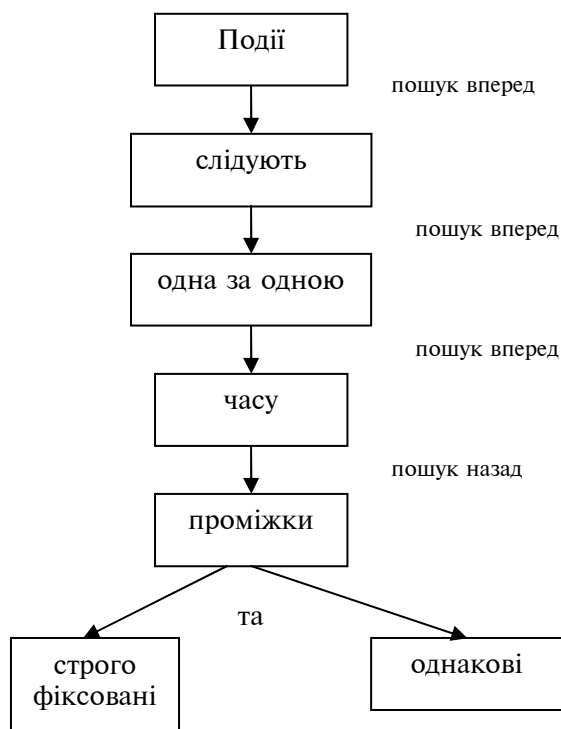


Рис.1. Приклад декомпозиції еталонної відповіді

Формування оцінки учня

Для формування оцінки завдання використовується апарат нечіткої математики. Застосування методів нечіткої математики дає можливість якісніше описувати характеристики як окремих тестових завдань, так і тестів у цілому, а також доступніше і якісно інтерпретувати результати тестування.

Завдання знаходження результуючого балу за тестове завдання здійснюється в кілька етапів: на першому етапі будується модель № 2 і модель № 3, системи нечіткого логічного висновку, для знаходження коефіцієнтів відповідностей першої та другої ключових лексем, що входять у відношення, із врахуванням важливості цих лексем. На другому етапі будується модель № 4 системи нечіткого логічного висновку для знаходження коефіцієнтів відповідностей семантичних відношень із врахуванням знайдених значень коефіцієнтів відповідностей першої та другої ключових лексем (отриманих при побудові моделі № 2 і моделі № 3), а та-

кож із врахуванням типу зв'язку (і, або, слідує за, слідує перед). Далі знайдені значення використовуються для побудови моделі № 1 для знаходження ступеня відповідності семантичного відношення еталонної відповіді і семантичного відношення з відповіді учня, з урахуванням важливості відношення, з урахуванням знайденого в моделі № 4 коефіцієнта відповідності семантичного відношення, а також із врахуванням наявності або відсутності в даному відношенні неприпустимої лексеми. Після цього будується модель № 5 нечіткого логічного висновку для знаходження набраного балу за завдання, що, зрештою, дасть змогу визначити, наскільки успішно учень відповів на поставлене запитання.

Опишемо кожну з моделей докладніше.

Модель № 1

Вхідні параметри:

- важливість семантичного відношення;
- коефіцієнт відповідності семантичного відношення;
- мітка наявності неприпустимої лексеми.

Вихідний параметр:

- ступінь відповідності семантичного відношення еталонної відповіді і семантичного відношення з відповіді учня (СВ).

Для знаходження значення лінгвістичної змінної “коефіцієнт відповідності семантичного відношення” (відношення, як було сказано вище, подано у вигляді $S_p(KS_p^{(1)}, KS_p^{(2)})$) необхідно додатково побудувати три моделі нечіткого логічного висновку.

Модель № 2

Вхідні параметри:

- коефіцієнт відповідності першої ключової лексеми, що входить у відношення, позначимо КВКЛ1;
- важливість першої ключової лексеми позначимо ВКЛ1.

Вихідний параметр:

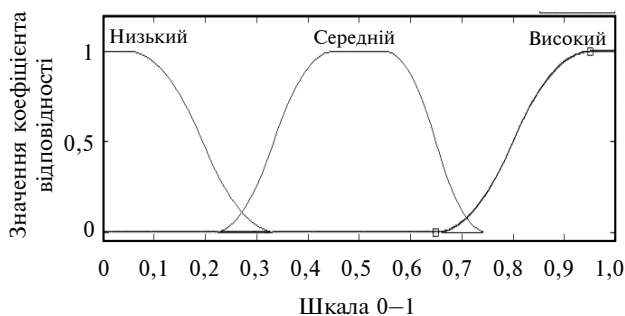
- коефіцієнт відповідності першої ключової лексеми, що входить у відношення (із врахуванням важливості); позначимо його КВКЛ1_В.

Згідно з правилами побудови нечітких моделей, опишемо множину лінгвістичних термів для лінгвістичної змінної “коефіцієнт відповідності першої ключової лексеми”, а також множину їх значень (табл. 1).

Функцію приналежності для КВКЛ1 зображено на рис. 2, а множину лінгвістичних термів для змінної ВКЛ1 наведено в табл. 2.

Таблиця 1. Множина лінгвістичних термів для КВКЛ1

Коефіцієнт відповідності першої ключової лексики	Шкала від 0 до 1
Низький	0–0,35
Середній	0,236–0,75
Високий	0,638–1

**Рис. 2.** Функції приналежності для КВКЛ1**Таблиця 2.** Множина лінгвістичних термів для ВКЛ1

Важливість першої ключової лексики	Шкала від 0 до 1
Низька	0–0,356
Середня	0,241–0,75
Висока	0,63–1

Множину лінгвістичних термів для КВКЛ1_В наведено в табл. 3, а функції приналежності для КВКЛ1_В – на рис. 3.

Таблиця 3. Множина лінгвістичних термів для КВКЛ1_В

Коефіцієнт відповідності першої ключової лексики (із врахуванням важливості)	Значення
Дуже низький	0–0,21
Низький	0,03–0,45
Середній	0,29–0,71
Високий	0,56–0,96
Дуже високий	0,79–1

Далі, для описаних змінних та лінгвістичних термів, що їм відповідають, будуються такі правила нечіткого логічного висновку.

1. Якщо “КВКЛ1”=низький і “ВКЛ1”=низька, то “КВКЛ1_В”=низький.

2. Якщо “КВКЛ1”=низький і “ВКЛ1”=середня, то “КВКЛ1_В”=низький.

3. Якщо “КВКЛ1”=низький і “ВКЛ1”=висока, то “КВКЛ1_В”=дуже низький.

4. Якщо “КВКЛ1”=середній і “ВКЛ1”=висока, то “КВКЛ1_В”=високий.

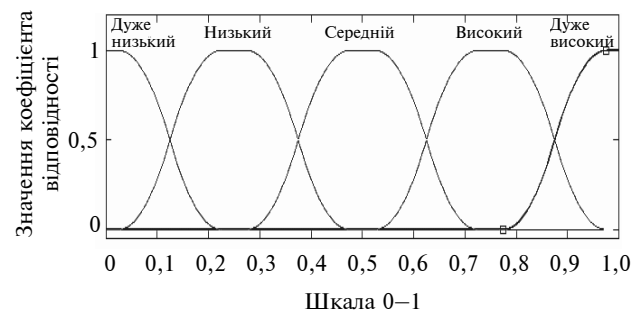
5. Якщо “КВКЛ1”=середній і “ВКЛ1”=середня, то “КВКЛ1_В”=середній.

6. Якщо “КВКЛ1”=середній і “ВКЛ1”=низька, то “КВКЛ1_В”=низький.

7. Якщо “КВКЛ1”=високий і “ВКЛ1”=низька, то “КВКЛ1_В”=середній.

8. Якщо “КВКЛ1”=високий і “ВКЛ1”=середня, то “КВКЛ1_В”=високий.

9. Якщо “КВКЛ1”=високий і “ВКЛ1”=висока, то “КВКЛ1_В”=високий.

**Рис. 3.** Функції приналежності для КВКЛ1_В

В результаті, припустимо, що лінгвістична змінна “коефіцієнт відповідності першої ключової лексики” дорівнює 0,6, що відповідає середньому значенню, “важливість першої ключової лексики” дорівнює 0,8, що відповідає високій важливості. Тоді згідно з алгоритмом Мамдани значення лінгвістичної змінної “коефіцієнт відповідності першої ключової лексики (із врахуванням важливості)” буде дорівнювати 0,75, що відповідає високому ступеню відповідності (рис. 4).

Модель № 3

Будується абсолютно аналогічно моделям № 1 і № 2.

Вихідні дані, отримані при побудові моделей № 2 і № 3, є вхідними даними для моделі № 4.

Модель № 4

Вхідні параметри:

- коефіцієнт відповідності першої ключової лексики, що входить у відношення (модель № 2);

- коефіцієнт відповідності другої ключової лексики, що входить у відношення (модель № 3);

- тип зв'язку.

Вихідний параметр:

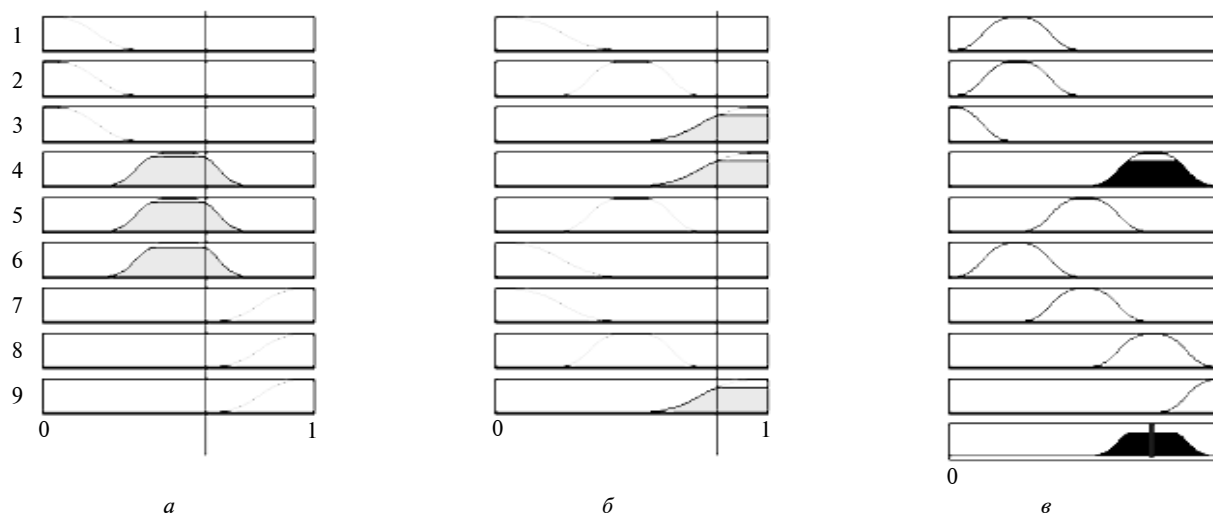


Рис. 4. Нечіткий логічний висновок за Мамдані для змінної КВКЛ1_В: *a* – коефіцієнт КЛ1 = 0,6; *б* – коефіцієнт КЛ1 = 0,814; *в* – коефіцієнт КЛ1 = 0,75

- коефіцієнт відповідності семантичного відношення.

Знаходження значень для змінних моделі № 4 відбувається за аналогічним алгоритмом, описаним у моделі № 2.

В результаті, знайдені за моделлю № 4 значення для вихідного параметра “коефіцієнт відповідності семантичного відношення” є значеннями для однойменного вхідного параметра для моделі № 1.

Лінгвістична змінна “мітка наявності неприпустимої ключової лексики” (див. модель № 1) є дискретною величиною, що набуває значення 0 або 1, де 0 означає, що зв’язка $S_p(KS_p^{(1)}, KS_p^{(2)})$ не містить неприпустимої лексики, 1 – містить. У разі, коли зв’язка $S_p(KS_p^{(1)}, KS_p^{(2)})$ містить неприпустиму лексему, вихідна лінгвістична змінна “ступінь відповідності семантичного відношення еталонної відповіді і семантичного відношення з відповіді учня” набуває нульового значення.

Таким чином, після проведення нечіткого логічного висновку і даних, отриманих при використанні моделей № 2–№ 4, формується множина значень, яка характеризує ступінь відповідності семантичних відношень еталонної відповіді і семантичних відношень з відповіді учня. Іншими словами, знаходиться множина значень для вихідної лінгвістичної змінної моделі № 1. В результаті відповідь учня можна подати у вигляді нечіткої множини (позначимо

її A), заданої на універсальній множині B еталонної відповіді, де $A = \{x, \mu_{a(x)}\}$ – x -зв’язка типу $S_p(KS_p^{(1)}, KS_p^{(2)})$ з відповіді учня; $\mu_{a(x)}$ – функція приналежності, що характеризує ступінь відповідності зв’язки x еталонному значенню. В результаті визначається ступінь рівності між двома множинами: еталонною множиною семантичних зв’язок та множиною семантичних зв’язок, отриманою при проведенні нечіткого логічного висновку. Отримані дані є значенням для лінгвістичної змінної “ступінь відповідності множини еталонної відповіді та множини учня”.

Розглянемо приклад. Нехай A – нечітка множина, задана на B , складається із значень вагових коефіцієнтів семантичних зв’язок; B – чітка множина, елементами якої є семантичні зв’язки з ваговим коефіцієнтом, що дорівнює 1:

$$A = \{0, 7; 0, 2; 0, 9; 0, 6; 0, 5; 1; 0, 4\}, \quad B = \{1, 1, 1, 1, 1, 1, 1\}.$$

Для розрахунку ступеня включення двох множин скористаємося імплікацією Лукашевича [9]. Маємо $V(A, B)$ і $V(B, A)$:

$$V(A, B) = (0, 7 \rightarrow 1) \& (0, 2 \rightarrow 0, 1) \& (0, 9 \rightarrow 1) \& (0, 6 \rightarrow 1) \& (0, 5 \rightarrow 1) \& (1 \rightarrow 1) \& (0, 4 \rightarrow 1) = 1;$$

$$V(B, A) = (1 \rightarrow 0, 7) \& (1 \rightarrow 0, 2) \& (1 \rightarrow 0, 9) \& (1 \rightarrow 0, 6) \& (1 \rightarrow 0, 5) \& (1 \rightarrow 1) \& (1 \rightarrow 0, 4) =$$

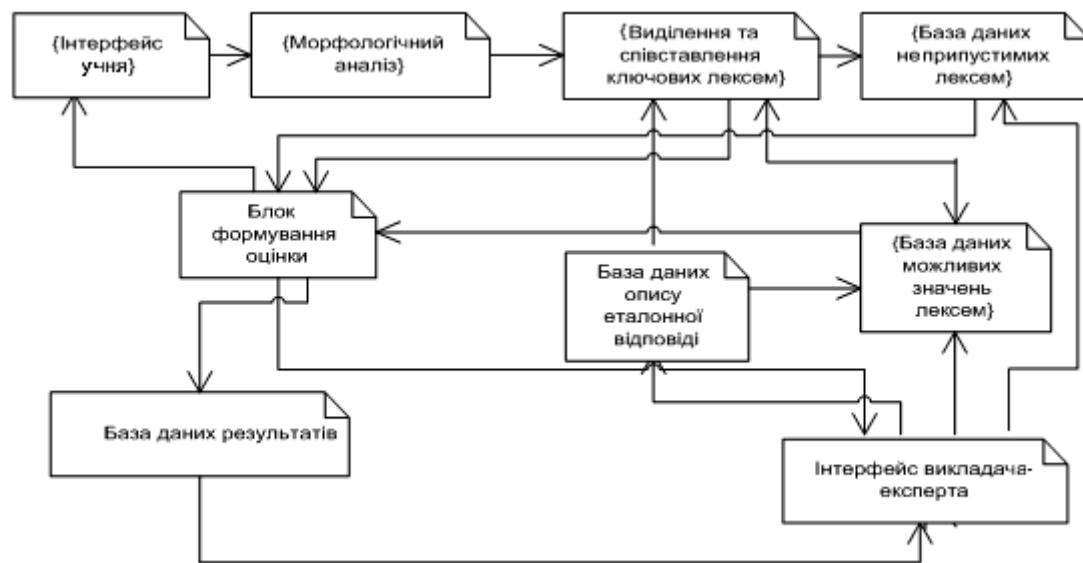


Рис. 5. Структурна схема системи тестування для тестів з відкритою формою тестового завдання

$$= (1 \& (1 - 1 + 0,7)) \& (1 \& (1 - 1 + 0,2)) \& (1 \& (1 - 1 + 0,9)) \& (1 \& (1 - 1 + 0,6)) \& (1 \& (1 - 1 + 0,5)) \& (1 \& (1 - 1 + 1)) \& (1 \& (1 - 1 + 0,4)) = 0,2.$$

Ступінь рівності між двома множинами: $\mu(A, B) = 0,2 \& 1$.

У разі, коли елементи нечіткої множини A близькі до 1, тобто множина A близька до еталонної множини, ступінь рівності між ними прямує до 1. Покажемо це на такому прикладі.

Нехай задано множини $A = \{1; 0,8; 0,9; 1; 1; 0,9; 1\}$, $B = \{1,1,1,1,1,1,1\}$, елементи множини A близькі до 1.

Аналогічно попереднім розрахункам маємо ступінь включення $V(B, A) = 0,8$ і $V(A, B) = 1$; $\mu(A, B) = 0,8 \& 0,1$.

І, нарешті, для знаходження результуючого балу за відповідь будується модель № 5 нечіткого логічного висновку (в дану модель входять результати, отримані на попередніх етапах при використанні моделей № 1–№ 4).

Модель № 5

Вхідні параметри:

- ступінь рівності множини еталонної відповіді та множини учня;
- складність завдання.

Вихідний параметр:

- бал за відповідь.

Множину лінгвістичних термів для лінгвістичної змінної “бал за відповідь” наведено в табл. 4.

Таблиця 4. Множина лінгвістичних термів для змінної “бал за відповідь”

Бал за відповідь	Значення
Відмінно	0,81–1
Добре	0,61–0,81
Задовільно	0,35–0,61
Незадовільно	0,12–0,35
Не допущено	0–0,12

Таким чином, після проведення нечіткого логічного висновку система визначає бал, який учень набрав при відповіді на запитання.

Структурну схему системи зображено на рис. 5.

Висновки

Система тестування знань, яка реалізує відкриту форму тестового завдання, на відміну від існуючих тестових систем забезпечує можливість формування відповіді учня на структурованій природній мові. Ця властивість системи істотно розширює її сферу застосування.

Система орієнтована на автоматизацію навчального процесу як для технічних, так і гуманітарних дисциплін, і може бути адаптована на будь-яку предметну область. Система може навчатися в процесі експлуатації.

Для подальшого вдосконалення системи необхідно розробити засоби для автоматизації створення формального опису еталонної відповіді згідно з налаштуваннями системи на конкретну предметну область.

Н.А. Соколова, П.А. Руденко, А.Т. Голуб

СИСТЕМА ТЕСТИРОВАНИЯ ЗНАНИЙ, РЕАЛИЗУЮЩАЯ ОТКРЫТУЮ ФОРМУ ТЕСТОВОГО ЗАДАНИЯ

Создана система тестирования, в основе которой лежат задания открытого типа. Эта задача является одной из наиболее известных задач в области искусственного интеллекта – задача машинного понимания текста на естественном языке. Система разработана для контроля знаний обучаемых при помощи методов нечеткой математики.

N.A. Sokolova, P.A. Rudenko, O.T. Golub

THE KNOWLEDGE TESTING SYSTEM WITH OPEN-TYPE TASKS

Based on open-type tasks, we develop the knowledge testing system. This task of machine understanding of texts in natural language is one of the most common in artificial intelligence. The developed system can be used for testing of gained knowledge by employing the methods of fuzzy mathematics.

1. Гіндін З.І., Леонтьєва Н.Н. Проблемы анализа и синтеза целого текста в системах машинного перевода, диалоговых та інформаційних системах. — М.: ВЦП, 1978.
2. Леоненко Л.Л. Алгоритмы оценки аналогичности текстов та їх застосування в комп'ютерному тестуванні. Система комп'ютерного тестування "Control" // VII Міжнар. конф. "Інтелектуальний аналіз інформації" IAI-2007, Київ, 15–18 травня 2007 р.— К., 2007.
3. Сулейманов Дж.Ш. Дослідження базових принципів побудови семантичного інтерпретатора текстів на природній мові. — http://ifets.ieee.org/russian/depository/v4_i3/html/2.html
4. <http://www.aot.ru/> Автоматична обробка тексту.
5. <http://testolog.narod.ru/Theory11.html> Теорія і методика педагогічних вимірювань.
6. Соколова Н.А., Руденко П.А. Побудова семантичного інтерпретатора відповідей навчаємих у вільно-конструйованій формі // VII Міжнар. конф. "Інтелектуальний аналіз інформації IAI-2007", Київ, 15–18 травня 2007 р. — К., 2007.
7. Соколова Н.А., Руденко П.А. Використання відкритої форми представлення тестового завдання для автоматизованого контролю знань // IV Міжнар. наук.-практ. конф. "Інформаційні технології та безпека в управлінні", Луганськ, 17 вересня 2007 р. — Луганськ, 2007.
8. Соколова Н.А., Руденко П.А. Співставлення відповіді навчуваного з еталонною при відкритій формі тестування // V Міжнар.наук.-практ. конф. "Інформаційні технології та безпека в управлінні", Луганськ, 15 вересня 2008 р. — Луганськ, 2008.
9. Малышев Н.Г. Нечеткие модели для экспертных систем в САПР. — М.: Энергоатом, 1991.— 20 с.

Рекомендована Радою
факультету прикладної математики
НТУУ "КПІ"

Надійшла до редакції
30 квітня 2009 року